

Szczecin, 16.02.2022

STRESZCZENIE ROZPRAWY DOKTORSKIEJ

Heterogeniczne oddziaływanie na procesy rozprzestrzeniania informacji w sieciach społecznych

Autor: mgr inż. Artur Karczmarczyk

Promotor: dr hab. inż. Jarosław Jankowski, prof. ZUT

Promotor pomocniczy: dr hab. inż. Jarosław Wątróbski, prof. US

Reprezentacja sieci społecznych w systemach elektronicznych ewoluowała od wczesnych systemów technicznych do zaawansowanych mediów społecznościowych integrujących mechanizmy komunikacji i interakcji podobne do tych znanych z realnego Świata. Rozwój platform społecznościowych wpłynął na potrzebę zrozumienia zachowań, wzorców i predyspozycji milionów użytkowników online i ich powiązania z zachowaniami w Świecie rzeczywistym.

Procesy rozprzestrzeniania informacji, obok relacji społecznych i aktywności online, należą do zjawisk absorbujących uwagę zarówno badaczy, jak i praktyków. W wielu przypadkach komunikacja elektroniczna, oparta na bazie procesów rozprzestrzeniania informacji w sieciach społecznych, daje wyniki lepsze niż tradycyjne kampanie reklamowe. Badania związane z dyfuzją treści cyfrowych zorientowane są wokół czynników wpływających na sukces kampanii, czynników wpływających na uczestnictwo użytkowników w procesie rozprzestrzeniania informacji, czy też wyboru użytkowników w sieci do inicjalizacji kampanii. Ponadto, badany jest wpływ roli różnych miar centralności podczas selekcji początkowych influencerów, role treści i struktur w sieciach, motywacja użytkowników do przekazywania treści, jak również rola emocji i innych czynników w procesie rozprzestrzeniania informacji.

Wiele wcześniejszych badań koncentrowało się na podejściach teoretycznych i empirycznych do maksymalizacji zasięgu, czyli zwiększaniu liczby węzłów, do których udało się dotrzeć w sieci. Chociaż zasięg, czyli liczba zainfekowanych węzłów w sieci, jest ważną miarą sukcesu kampanii, z praktycznego punktu widzenia kampanie rozprzestrzeniania informacji w sieciach społecznych mogą mieć różne cele i specyfikę. Inna strategia może być wykorzystana w celu pozyskania dużej liczby potencjalnych odbiorców w bardzo krótkim czasie niż w przypadku potrzeby osiągnięcia organicznego wzrostu bazy odbiorców o zadanych cechach. Aby uwzględnić różne cele, można wykorzystać wielokryterialną ocenę procesu i dobrać parametry oraz cele zgodnie z preferencjami i priorytetami.

Głównym celem rozprawy doktorskiej było opracowanie i weryfikacja algorytmów heterogenicznego oddziaływania na procesy propagacji informacji w sieciach złożonych z udziałem złożonych rankingów, z uwzględnieniem wielokryterialnej oceny efektywności, w celu zwiększenia zasięgu procesu, jego dynamiki lub innych charakterystyk zgodnych z preferencjami decydenta.

Praca powstała jako cykl 10 powiązanych tematycznie artykułów – cztery artykuły opublikowane w międzynarodowych czasopismach z otwartym dostępem, pięć artykułów wydanych w recenzowanych materiałach konferencyjnych oraz jeden rozdział monografii. Suma punktów ministerialnych cyklu wynosi 875 (suma ważona 518). Sumaryczny współczynnik Impact Factor w cyklu to 10.762.

16.02.2022

Artur Karczmarczyk

Szczecin, 16.02.2022

DOCTORAL THESIS ABSTRACT

Heterogeneous influence on the information diffusion processes in social networks

Author: mgr inż. Artur Karczmarczyk

Supervisor: dr hab. inż. Jarosław Jankowski, prof. ZUT

Auxiliary supervisor: dr hab. inż. Jarosław Wątróbski, prof. US

The representation of social networks in electronic systems has evolved from early technical systems to advanced social media integrating communication and interaction mechanisms similar to those known from the real world. The development of social media platforms has influenced the need to understand the behaviors, patterns and predispositions of millions of online users and how they relate to behaviors in the real world.

The processes of diffusing information, apart from social relations and online activity, are phenomena that attract the attention of both researchers and practitioners. In many cases, electronic communication, based on the processes of diffusing information in social networks, gives better results than traditional advertising campaigns. Research related to the diffusion of digital content is focused on factors influencing the success of a campaign, factors influencing the participation of users in the process of disseminating information, or the choice of users on the Internet to initiate a campaign. Moreover, the influence of the role of various centrality measures in the selection of initial influencers, the roles of content and structures in networks, user motivation to communicate content, as well as the role of emotions and other factors in the information diffusion process are investigated.

Much previous research has focused on theoretical and empirical approaches to maximizing coverage, i.e. increasing the number of nodes reached in the network. While coverage is an important measure of campaign success, from a practical point of view, information diffusion campaigns on social networks can have different goals and specificities. Different strategy can be used to acquire a large number of potential customers in a very short time than in the case of the need to achieve organic growth of the customer base with given characteristics. In order to take into account different goals, one can use multi-criteria evaluation of the process and select parameters and goals according to the decision-maker's preferences and priorities.

The main goal of the doctoral dissertation was to develop and verify algorithms for heterogeneous influence on the processes of information propagation in complex networks with the participation of complex rankings, taking into account multi-criteria assessment of effectiveness, in order to increase the coverage of the process, its dynamics or other characteristics according to the preferences of the decision-maker.

The dissertation was compiled as a series of 10 thematically related articles – four articles published in international open-access journals, five articles published in peer-reviewed conference materials and one chapter of a monograph. The sum of ministerial points for the cycle is 875 (weighted sum 518). The total Impact Factor of the cycle is 10.762.

16.02.2022.

Artur Karczmarczyk

Dr hab. inż. Dariusz Król, prof. uczelni
Katedra Informatyki Stosowanej
Wydział Informatyki i Telekomunikacji
Politechnika Wrocławska
E-mail: Dariusz.Krol@pwr.edu.pl

Recenzja rozprawy doktorskiej
mgra inż. Artura Karczmarczyka
pt. „Heterogeniczne oddziaływanie
na procesy rozprzestrzeniania informacji w sieciach społecznych”

Niniejsza recenzja została przygotowana w odpowiedzi na pismo Prorektora ds. Nauki Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie prof. dr. hab. inż. Jacka Przepiórskiego z dnia 8 czerwca 2021 roku.

1. Kompozycja i treść rozprawy

Recenzowana rozprawa mgra inż. Artura Karczmarczyka pt. „Heterogeniczne oddziaływanie na procesy rozprzestrzeniania informacji w sieciach społecznych” została przygotowana pod kierunkiem promotora dra hab. inż. Jarosława Jankowskiego, profesora Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie oraz promotora pomocniczego dra hab. inż. Jarosława Wątróbskiego, profesora Uniwersytetu Szczecińskiego. Obejmuje 212 stron tekstu i składa się z dwóch zasadniczych części: streszczenia rozprawy doktorskiej oraz pełnych tekstów opublikowanych prac naukowych ponumerowanych od A1 do A10.

W części pierwszej znajduje się syntetyczny opis uzyskanych wyników w postaci przedstawienia problemu badawczego i głównego celu rozprawy, omówienia cyklu publikacji wchodzących w skład rozprawy, opisanie sposobów oddziaływania na procesy rozprzestrzeniania informacji w sieciach społecznych oraz środowiska symulacyjnego do badania procesu dyfuzji informacji w



sieciach złożonych, podsumowania, wreszcie zreferowania dorobku naukowego, dydaktycznego, organizacyjnego i zawodowego kandydata. Opis ten uzupełnia spis rysunków, spis tabel oraz spis literatury zawierający 42 pozycje. Część druga rozprawy zawiera dziesięć współautorskich prac naukowych opublikowanych w latach 2018-2021, w których przedstawione są szczegółowe wyniki badań związanych z realizacją celów badawczych pracy.

Rozprawa koncentruje się na problemach wpływania na zasięg, dynamikę i inne charakterystyki procesu rozprzestrzeniania informacji. Najbardziej istotnymi opracowanymi rozwiązaniami są: a) wielokryterialne planowanie i ewaluacja procesów rozprzestrzeniania informacji w sieciach społecznych, b) oddziaływanie na inicjalizację procesów rozprzestrzeniania z udziałem próbkowania i model doboru wielkości próbek w ujęciu wielokryterialnym, c) oddziaływanie na proces propagacji informacji poprzez wielokryterialny dobór rankingów dla węzłów zasiewowych w podejściu sekwencyjnym, d) oddziaływanie poprzez nierównomierny rozrzut prawdopodobieństwa propagacji informacji, e) oddziaływanie poprzez wielokryterialne targetowania w wieloatrybutowe węzły. Pracę uzupełnia opracowanie dotyczące otwartego zorientowanego obiektowo środowiska symulacyjne do badania procesu dyfuzji informacji w sieciach złożonych.

Oceniając tytuł rozprawy, tj. „Heterogeniczne oddziaływanie na procesy rozprzestrzeniania informacji w sieciach społecznych”, należy stwierdzić, iż został on sformułowany we właściwy sposób, a zakres merytoryczny zaprezentowanych analiz i badań w pełni odpowiada tytułowi.

Teza rozprawy została sformułowana poprawnie, choć pozostawiła pewien niedosyt w zakresie jej uzasadnienia. Poza krótkim opisem problemu badawczego zawartym w rozdz. 1 brakuje szerszego wyjaśnienia, z jakiego powodu Autor zdecydował się na takie, a nie inne rozwiązanie. W tym kontekście istotny niedostatek stanowi brak podsumowań kolejnych rozdziałów, które mogłyby referować krytyczną analizę oraz przedstawienie argumentów stojących za przyjętym sposobem prowadzenia badań. Podsumowanie ujęte w rozdz. 6 zawiera listę najważniejszych osiągnięć. Brakuje natomiast podsumowującej dyskusji na temat ograniczeń zaproponowanych rozwiązań oraz szerszego spojrzenia na potencjalne obszary dalszych badań i zastosowań.



Ponieważ rozprawa bazuje na zbiorze powiązanych tematycznie wieloautorskich prac naukowych, nieodłącznym jej elementem powinien być załącznik zawierający oświadczenia współautorów poszczególnych prac o ich wkładzie poprzez procentowe określenie udziału i wskazanie na czym polegał indywidualny wkład każdej z osób. Autor rozprawy podał wprawdzie obok udziału procentowego także krótki opis swojego wkładu w powstaniu każdej z dziesięciu prac, jednakże nie jest to oświadczenie wszystkich współautorów. Moim zdaniem, jeżeli dany artykuł wieloautorski ma być włączony do jednotematycznego zbioru artykułów składającego się na pracę doktorską jednego ze współautorów, to pozostali współautorzy powinni być poinformowani o tym zamiarze i powinni wyrazić na to zgodę. Uwaga ta ma charakter techniczny i może być uwzględniona bez naruszania treści rozprawy, np. poprzez przedstawienie odpowiednich oświadczeń współautorów w ramach dyskusji nad rozprawą.

2. Problem badawczy i jego znaczenie

Tematyka rozprawy obejmuje badania nad procesami rozprzestrzeniania informacji w sieciach społecznych. Głównym problemem podjętym w pracy jest heterogeniczne oddziaływanie na te procesy. Definicja heterogeniczności oddziaływania podana przez Autora nie jest formalna, więc może nieść wątpliwości. W dużym skrócie oznacza różne sposoby wpływania na zasięg, dynamikę i inne charakterystyki procesu rozprzestrzeniania informacji m.in. poprzez dobór węzłów początkowych, nierównomierny rozrzut prawdopodobieństwa procesu propagacji, czy zróżnicowane sekwencje inicjalizacyjne.

Problematyka ta uważana jest za jeden z ważniejszych obszarów i kierunków rozwojowych obliczeniowej nauki o sieciach społecznych, która łączy informatykę z jej głębokim rozumieniem danych rzeczywistych, z nauką o dużych zbiorach danych. Te i jeszcze inne argumenty za tym, iż oceniana praca dotyczy ważnych problemów badawczych w informatyce, znajdują się przede wszystkim w rozdziale pierwszym i czwartym.

W dużym uproszczeniu można powiedzieć, że recenzowana rozprawa ma charakter przede wszystkim konstrukcyjno-symulacyjny. Problem, któremu poświęcona jest rozprawa, może też mieć duże znaczenie praktyczne. Możliwość opracowania algorytmów heterogenicznego



oddziaływania na procesy propagacji informacji w sieciach społecznych z udziałem złożonych rankingów dynamicznych, mechanizmów topologicznych oraz z uwzględnieniem wielokryterialnej oceny efektywności może mieć bardzo duże znaczenie także w wielu innych działach informatyki, takich jak: bezpieczeństwo komputerowe, inżynieria oprogramowania, inżynieria wiedzy, symulacja komputerowa, komunikacja sieciowa oraz systemy wspomaganie decyzji.

Uważam, że opiniowana rozprawa podejmuje ważny i aktualny problem. Podjęcie przedstawionej problematyki jest uzasadnione zarówno ze względów poznawczych, jak i potencjalnych możliwości wielu praktycznych zastosowań związanych ze spełnianiem ciągle rosnących oczekiwań użytkowników współczesnych systemów komputerowych.

3. Oryginalne osiągnięcia

Do najważniejszych osiągnięć rozprawy, decydującymi o wartości teoretycznej i praktycznej, wyróżniających je spośród dostępnych w literaturze przedmiotu, można zaliczyć:

- a) **opracowanie podejścia do wielokryterialnego planowania i ewaluacji procesów rozprzestrzeniania informacji w sieciach społecznych.** W rozdziale 4.2 odnoszącym się do publikacji A1 i A2 wymieniono strategie optymalizujące trzy kryteria: liczbę węzłów zasiewowych, prawdopodobieństwo propagacji informacji i potencjalny zasięg, jaki można uzyskać. Wykazano, że najlepsza strategia zależy od preferencji decydenta. Przeprowadzone badania empiryczne zostały oparte na sieci rzeczywistej o 7610 węzłach i 15751 krawędziach. W ramach badań wygenerowano 150 sieci syntetycznych. Następnie przeprowadzono łącznie 4000 symulacji dla 400 zestawów kryteriów typu liczba węzłów początkowych versus miara centralności stopnia lub wektora własnego. W celu jak najlepszego dopasowania sieci syntetycznej do docelowej sieci rzeczywistej wykorzystano miarę dywergencji Kullbacka-Leiblera (KLD) oraz analizę wielokryterialną. W efekcie uzyskano czas dotarcia do założonej liczby użytkowników reprezentowany przez liczbę kroków symulacji względem całkowitej liczby użytkowników sieci.



- b) **opracowanie podejścia do wyboru rozmiaru próbek sieci rzeczywistej w celu uproszczonego obliczeniowo pozyskiwania przyszłych rekomendacji odnośnie parametrów inicjalizacji procesów w sieci rzeczywistej.** W rozdziale 4.3 odnoszącym się do publikacji A3 i A4 zaproponowano wykorzystanie próbek sieci rzeczywistej. Badanie empiryczne oparto na sieci rzeczywistej Gnutella. Odwzorowana sieć składa się z 8846 węzłów i 31839 krawędzi. W badaniu wykorzystano taki sam jak poprzednio zestaw kryteriów ewaluacyjnych strategii: liczbę węzłów zasiewowych oraz zasięg sieci. Planowanie procesu na sieci rzeczywistej w oparciu o jej próbki przeprowadzono z wykorzystaniem trzech sieci różnej wielkości: 10%, 30% i 50% liczby węzłów oryginalnej sieci. Próbki uzyskano z wykorzystaniem metody *snowball sampling*. Wyniki najbardziej zbliżone do sieci rzeczywistej uzyskano dla próbek o jak najwyższej liczbie węzłów. Kryteria wyboru próbki sieci rzeczywistej – wielkość, czas generowania, zasięg oraz czas trwania symulacji – pogrupowano według kosztów i dokładności. Na podstawie analizy wizualnej GAIA dla zwykłej funkcji preferencji wskazano zwycięską próbkę, a po zastosowaniu liniowej funkcji preferencji można było wskazać także drugi i kolejny wybór.
- c) **opracowanie podejścia wykorzystującego sieci syntetyczne i zróżnicowanie sekwencji inicjalizacji węzłów zasiewowych w celu doboru strategii do inicjalizacji i realizacji procesu w sieci rzeczywistej.** W rozdziale 4.4 odnoszącym się do publikacji A5 i A10 poddano analizie oddziaływanie na proces propagacji informacji poprzez wielokryterialny dobór rankingów dla węzłów początkowych w podejściu sekwencyjnym. Zaproponowano rozszerzenie wielokryterialnego podejścia o sekwencyjną inicjalizację węzłów z wykorzystaniem symulacji w sieciach syntetycznych. W tym przypadku wykorzystano nowy zestaw parametrów: odsetek węzłów, prawdopodobieństwo propagacji, liczba iteracji inicjalizacji procesu, interwał między sekwencjami zasilania sieci, miara centralności oraz dwa kryteria wydajnościowe: liczba iteracji i uzyskany zasięg. Do planowania wykorzystano 15 sieci syntetycznych o wielkości 50% sieci Gnutella. Na podstawie miary KLD do dalszych badań wybrano jedną z sieci opartą na modelu bezskalowym (BA). Dalej analizowano wpływ liczby iteracji inicjalizacji procesu oraz interwał między sekwencjami zasilania sieci. Wraz ze wzrostem liczby iteracji lub



interwału między sekwencjami zaobserwowano logarytmiczny wzrost średniego pokrycia oraz liniowy wzrost czasu trwania symulacji.

- d) **zbadanie oddziaływania nierównomiernego rozrzutu prawdopodobieństwa propagacji informacji w sieciach na osiągnięty zasięg i dynamikę procesu rozprzestrzeniania informacji.** W rozdziale 4.5 odnoszącym się do publikacji A6 przedstawiono oddziaływanie wpływu na dynamikę i zasięg procesu dyfuzji informacji poprzez nierównomierny rozrzut prawdopodobieństwa propagacji. Zauważano, że zwiększanie motywacji użytkowników o niskich wartościach miar centralności może być wydajniejsze pod względem kosztów, w związku z potencjalnie mniejszymi zachętami niż dla popularnych węzłów centralnych. W trakcie badań opracowano szereg algorytmów generowania wektorów rozkładów prawdopodobieństw dla zadanej liczby węzłów i średniego prawdopodobieństwa propagacji, m.in. jednostajny, proporcjonalny, odwrotnie proporcjonalny, normalny, geometryczny i odwrotnie geometryczny. Na podstawie symulacji największym wzrostem zasięgu skutkowały podejścia oparte na odwróconym rozkładzie geometrycznym i odwróconym rozkładzie proporcjonalnym. W pozostałych przypadkach zaobserwowano zmniejszenie zasięgu.
- e) **zaproponowanie podejścia do wielokryterialnego targetowania w wieloatrybutowe węzły w sieciach społecznych.** W rozdziale 4.6 odnoszącym się do publikacji A7, A8 i A9 Autor zajmuje się problemem zasięgu osiągniętego bezpośrednio w grupie docelowej przy jednoczesnym unikaniu efektu habituacji. W proponowanym podejściu rozważano superpozycję grupy atrybutów, np. płeć i wiek oraz stopnia węzła jako złożone kryterium decyzyjne w celu wybrania węzłów początkowych. To umożliwiło uzyskanie w grupie docelowej zasięgu większego niż w przypadku tradycyjnego podejścia opartego o miarę centralności stopnia węzła przy niewielkim zmniejszeniu zasięgu globalnego.
- f) **opracowanie otwartego zorientowanego obiektowo środowiska symulacyjnego do badania procesu dyfuzji informacji w sieciach złożonych.** W rozdziale 5 odnoszącym się do publikacji A10 opisano otwarte zorientowane obiektowo środowisko symulacyjne do badania procesu dyfuzji informacji. Środowisko umożliwia przeprowadzać eksperymenty



heterogenicznych scenariuszy oddziaływania na procesy rozprzestrzeniania informacji, takich jak:

- oddziaływanie przez zmienną frakcję węzłów zasiewowych;
- oddziaływanie przez wysiewanie pojedyncze lub sekwencyjne w wielu iteracjach;
- oddziaływanie przez nierównomierny rozrzut prawdopodobieństwa propagacji;
- ewaluacja i planowanie marketingu wirusowego w sieciach społecznych w oparciu o sparametryzowane wartości frakcji wysiewu, prawdopodobieństwo propagacji, miar centralności węzłów i rankingi.

Zgodnie z moją wiedzą w literaturze przedmiotu nie raportowano wcześniej rozwiązań tak uniwersalnych jak propozycje a)-f) przedstawione w rozprawie. Także przyszłe prace badawcze, np. nad nowymi metodami oddziaływania na procesy rozprzestrzeniania informacji, mogą zostać z łatwością umieszczone w kontekście już opracowanych metod.

Załączona bibliografia zawiera w zdecydowanej większości prace aktualne, opublikowane w przeciągu ostatnich dwóch dekad. Dzięki temu czytelnik może łatwo ocenić wkład Autora do omawianej dyscypliny.

W moim przekonaniu uważam, że Autor w pełni zrealizował założony cel rozprawy i tym samym wykazał słuszność postawionej tezy. Uzyskane rezultaty potwierdzają jego kompetencje z zakresu informatyki technicznej i nauki o sieciach. Dowodzą również, że Autor potrafi podejmować i samodzielnie realizować zaplanowane cele badawcze.

4. Poprawność

Badania zostały poprawnie zaplanowane i przeprowadzone. W dostatecznym stopniu, choć z dużym niedosytem, przedstawiono w pracy: problem badawczy (rozdział pierwszy), główny cel rozprawy (rozdział drugi), metody oddziaływania na procesy propagacji informacji w sieciach społecznych (rozdział czwarty), środowisko symulacyjne do badania procesu dyfuzji informacji w sieciach złożonych (rozdział piąty) oraz podsumowanie (rozdział szósty). W mojej opinii Autor dotrzymał staranności w opracowaniu metod, dzięki czemu wyniki pięciu eksperymentów (rozdziały 4.2-4.6) są wiarygodne i interesujące.



Wśród niedostatków rozprawy w pierwszej kolejności należy wymienić brak formalnych modeli oddziaływania na procesy rozprzestrzeniania informacji w sieciach społecznych oraz brak analizy złożoności obliczeniowej proponowanych algorytmów generacji wektora dla różnych rozkładów. Z tego względu trudno jest ocenić możliwość wykorzystania proponowanego podejścia w bardziej złożonych przypadkach lub przypadkach o innej charakterystyce dystrybucji. Czasami brakuje również porównania analizowanego rozwiązania z podejściami opisywanymi w literaturze. Tego typu badania pozwoliłyby dokładniej określić obszar zastosowań opracowanych podejść.

Lektura rozprawy skłania do kilku uwag o charakterze polemicznym. Poniżej zamieszczam listę trzech zauważonych przeze mnie punktów niejasnych lub dyskusyjnych.

- Zasadniczo, praca posiada jedną poważną wadę, którą stanowi brak próby przedstawienia analitycznego podejścia do problemu oddziaływania na procesy rozprzestrzeniania informacji. Rozprawa, jak już pisałem, ma charakter konstrukcyjno-symulacyjny i składa się z dziesięciu publikacji A1-A10 opisanych w pierwszej części rozprawy. Brakuje zdefiniowania modelu matematycznego, który, dla danego zbioru kryteriów, przewidywałby zachowanie się metod oddziaływania na poszczególnych etapach. Autor nie podjął próby przeprowadzenia rozważań teoretycznych. Wydaje się, że istnieją pewne formalizmy, np. dla wprowadzania zmian, które można by spróbować dostosować do specyfiki opisywanych metod. Gwoli ścisłości i na obronę Autora dodać należy, że sformułowanie takiego modelu jest zadaniem trudnym ze względu na złożony i bardzo zróżnicowany charakter wybranych obszarów oddziaływania.
- Drugą, ale mniej istotną wadą jest sposób przedstawienia cyklu publikacji A1-A10 w pierwszej części rozprawy, często mało czytelny i niespójny, bez analizy krytycznej literatury, w szczególności całościowego systematycznego przeglądu, np. zgodnie ze schematem PRISMA. Zabrakło tutaj pełnej analizy możliwych czynników oddziaływania, np. określenie kompletnego łańcucha przyczynowego, a nie tylko wybranych czynników z załączonych publikacji cyklu. A przecież w tezie pracy jest mowa o zróżnicowanych metodach oraz o innych charakterystykach procesu zgodnych z preferencjami decydenta. Zastrzeżenia budzi także zestawianie, podsumowanie i raportowanie wyników, np. analiza jakościowa, z której wnioski byłby ujęte tematycznie.



- Ostatnią widoczną wadą jest brak analizy kluczowych ograniczeń przyjętych w badaniach i potencjalnych rozwiązań alternatywnych, co w rezultacie w pewnych warunkach może spowodować nawet zmniejszenie zasięgu procesu rozprzestrzeniania informacji. W badaniach pominięto problem możliwego silnego skorelowania wykorzystywanych parametrów oraz problem złożoności opracowanych algorytmów heterogenicznego oddziaływania. Końcowym krokiem powinno być odniesienie uzyskanych wyników jako wytycznych do dalszych badań. Z reguły rezultaty analizy ograniczeń bezpośrednio służą do sformułowania rekomendacji tematów przyszłych badań i do zidentyfikowania luk badawczych na podstawie wariantowych scenariuszy. Takie scenariusze stosowane są jako narzędzia poszukiwawcze oraz narzędzia decyzyjne ukazujące możliwe wybory i ich potencjalne konsekwencje.

Pod względem redakcyjnym należy zaznaczyć, że rozprawa została przygotowana poprawnie, posiada właściwą strukturę i proporcje. Poziom edycyjny tekst jest dobry. Biorąc pod uwagę objętość pracy, liczba pomyłek jest mała i nie umniejsza wartości pracy.

W skrócie do zalet i mocnych stron rozprawy zaliczam: opracowanie pięciu podejść do inicjalizacji i realizacji procesu rozprzestrzeniania informacji, przeprowadzenie rozległych i skrupulatnych badań eksperymentalnych, opracowanie otwartego zorientowanego obiektowo środowiska symulacyjnego do badania procesu dyfuzji informacji w sieciach. Równocześnie dostrzegam następujące wady i słabe strony rozprawy: brak próby przedstawienia analitycznego rozwiązania problemu, mało czytelny i niespójny sposób przedstawienia cyklu publikacji, brak analizy kluczowych ograniczeń i rozwiązań alternatywnych. Uważam, że pomimo szeregu, sformułowanych powyżej uwag krytycznych opiniowana rozprawa jest próbą, chociaż nie do końca udaną, oryginalnego podejścia do zagadnienia heterogenicznego oddziaływania na procesy rozprzestrzeniania informacji w sieciach społecznych. Recenzowana rozprawa spełnia w sposób zaledwie dostateczny warunki i wymagania odpowiedniej ustawy o tytule i stopniach naukowych w odniesieniu do rozpraw doktorskich.



5. Dorobek naukowy kandydata do stopnia doktora

Zbiór dziesięciu prac naukowych załączonych do rozprawy odzwierciedla wysoki poziom wiedzy i doświadczenia praktycznego Autora w przedmiotowej tematyce badań. Bogata i umiejętnie cytowana w tych pracach literatura robi bardzo dobre wrażenie. Badania prowadzone w ramach rozprawy były częściowo realizowane w ramach grantów NCN w konkursie OPUS 11 numer 2016/21/B/HS4/01562 pt. „Wspomaganie procesów rozprzestrzeniania treści marketingowych w mediach społecznościowych” oraz Opus 14 numer 2017/27/B/HS4/01216 pt. „Nieinwazyjne metody redukcji zjawiska habituacji w marketingu elektronicznym”, w których Autor rozprawy był wykonawcą.

Z załączonego zbioru publikacji należy wyróżnić pracę A2 nie tylko ze względu na wysoki współczynnik wpływu $IF= 5.341$, ale także ze względu na uzyskaną liczbę cytowań według bazy Scopus równą 168. Sumaryczne wyniki współczynnika wpływu dla prac A1-A10 o wartości powyżej 10, czy dla wszystkich publikacji Autora o wartości powyżej 19, świadczą o wysokim walorze tych prac. Nie ulega żadnej wątpliwości, że te wskaźniki bibliometryczne można ocenić jako bardzo dobre dla młodego naukowca.

O wiedzy i umiejętnościach Autora świadczą nie tylko uzyskane wyniki w rozprawie, ale także jego pozostałe osiągnięcia publikacyjne, dydaktyczne, organizacyjne i zawodowe. Mgr inż. Artur Karczmarczyk jest współautorem 30 prac naukowych zarejestrowanych w bazie Scopus ze wskaźnikiem Hirscha $h=12$. Wymienione publikacje były dotychczas cytowane 588 razy. Tematyka prac głównie koncentruje się wokół zagadnień metodycznych i praktycznych związanych z wykorzystaniem oraz rozwojem metod MCDA, ale także dotyczy metod oceny jakości serwisów internetowych oraz metod wspomaganie procesów decyzyjnych w obszarze problemów zrównoważonego rozwoju.

Na szczególne podkreślenie zasługuje otrzymanie przez kandydata nagrody Best Paper Award za publikację pt. „Integrated approach to e-commerce websites evaluation with the use of surveys and eye tracking based experiments” na konferencji FedCSIS 2017.



Biorąc pod uwagę całokształt dorobku badawczego i publikacyjnego mgr inż. Artura Karczmarczyka stwierdzam, że kandydat posiada znaczny ilościowo i istotny udokumentowany jakościowo dorobek naukowy.

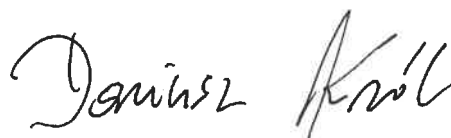
6. Wnioski końcowe

Mgr inż. Artur Karczmarczyk wykazał się znajomością i umiejętnością analizy literatury naukowej przedmiotu rozprawy, konstruowania złożonych algorytmów i eksperymentów symulacyjnych, a także zaplanowania i przeprowadzania poprawnej weryfikacji uzyskanych wyników.

Uważam, że:

- **rozprawa zawiera oryginalne rozwiązanie problemu naukowego,**
- **kandydat posiada ogólną wiedzę teoretyczną w dyscyplinie informatyka,**
- **kandydat posiada umiejętność samodzielnego prowadzenia pracy naukowej, w szczególności modelowania, optymalizacji i prowadzenia eksperymentów symulacyjnych.**

Moim zdaniem opiniowana rozprawa w stopniu wystarczającym spełnia warunki stawiane przez obowiązującą ustawę o stopniach i tytule naukowym w odniesieniu do rozpraw doktorskich w dyscyplinie informatyka techniczna i telekomunikacja w myśl art. 13 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym (Dz. U. Nr 65 z 2003 r., poz. 595 z późn. zm.). Wnoszę o jej przyjęcie oraz dopuszczenie do publicznej obrony.





POLISH-JAPANESE ACADEMY OF INFORMATION TECHNOLOGY

ポ
ー
ラ
ン
ド
日
本
情
報
工
科
大
学

Warszawa, 30.08.2021

prof. dr hab. Adam Wierzbicki
Polsko-Japońska Akademia
Technik Komputerowych

OPINIA O ROZPRAWIE DOKTORSKIEJ

MGR ARTURA KARCZMARCZYKA

“HETEROGENICZNE ODDZIAŁYWANIE NA PROCESY ROZPRZESTRZENIANIA INFORMACJI W SIECIACH SPOŁECZNYCH”

1. Jakie zagadnienie naukowe jest rozpatrzone w pracy i czy zostało ono dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?

Rozprawa poświęcona jest metodom planowania oddziaływania na procesy rozprzestrzeniania informacji w sieciach społecznych. Oddziaływanie na te procesy może polegać na wyborze węzłów początkowych (*seeds*), które jednak nie muszą być wybrane jednorazowo na początku rozprzestrzeniania informacji – możliwe jest dobieranie węzłów w ustalonej kolejności (*sequential seeding*). Inną możliwością oddziaływania na rozprzestrzelenie informacji jest oferowanie zachęt uczestnikom sieci społecznych, które mogą ich skłonić do przekazania informacji swoim sąsiadom w sieci. Te różnorodne możliwości oddziaływania określane są w pracy jako heterogeniczne.

Problem planowania oddziaływania na procesy rozprzestrzeniania informacji jest przedstawiony w rozprawie jako problem wielokryterialny. Kryteriami używane przez decydenta do planowania procesu mogą być: zasięg (liczba uczestników sieci społecznej, do których dotrze rozprzestrzeniająca się informacja), dynamika rozprzestrzeniania się informacji, koszt (poniesiony za wynagradzanie



POLISH-JAPANESE ACADEMY OF INFORMATION TECHNOLOGY

ポ
ー
ラ
ン
ド
日
本
情
報
工
科
大
学

wybranych uczestników do pierwszego przesłania informacji lub za zachęcanie uczestników do rozprzestrzeniania informacji). Można także wyobrazić sobie inne kryteria, na przykład rozprzestrzenienie się informacji do wyselekcjonowanych węzłów.

Rozprawa ma charakter teoretyczny.

2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł (w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle) świadczącej o dostatecznej wiedzy autora? Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?

Rozprawa jest zbiorem publikacji, poprzedzonym obszernym wstępem. Każda z publikacji ma własną bibliografię; jednak bibliografia wstępu (a zatem i całej rozprawy) ma zaledwie 42 pozycje. Uważam, że nie jest to właściwy sposób redakcji rozprawy złożonej z publikacji. Według mnie, rozprawa taka powinna być poprzedzona wstępem, który stanowi wyczerpujące i niezależne od innych publikacji wprowadzenie do tematyki rozprawy. Dlatego przegląd literatury pozostawia istotny niedosyt.

Co więcej, doktorant pomija (wykazując nieznaną tematu) istotne metody wspomaganie decyzji wielokryterialnych. W rozprawie rozpatrywane są wyłącznie metody wspomaganie decyzji wyrażające preferencje użytkowników w formie wag, których ustalenie wymaga porównań parami. Nie

jest to jedyna ani najlepsza metoda. Wśród alternatywnych metod najistotniejszą jest metoda punktu odniesienia, która pozwala decydentowi wyrażać preferencje dla każdego z kryteriów oddzielnie (niezależnie od innych kryteriów), oraz która pozwala na wyrażenie preferencji, które mogą doprowadzić do wyboru dowolnego rozwiązania Pareto- optymalnego. Własności tej nie posiadają metody oparte na sumie ważonej jako funkcji celu, albowiem wagi pozwalają na wybranie wyłącznie rozwiązań Pareto- optymalnych znajdujących się na otoczce wypukłej zbioru dostępnych rozwiązań. Proponuję lekturę następujących pozycji oraz uzupełnienie bibliografii o te pozycje:

1. Dujardin, Y., & Chadès, I. (2018). Solving multi-objective optimization problems in conservation with the reference point method. *PLoS one*, 13(1), e0190748.
2. Romero, C., Tamiz, M., & Jones, D. F. (1998). Goal programming, compromise programming and reference point method formulations: linkages and utility interpretations. *Journal of the Operational Research Society*, 49(9), 986-991.
3. Wierzbicki, A. P. (1998). Reference point methods in vector optimization and decision support.

3. Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Autor użył jako głównej metody badawczej symulacji komputerowej. Opracował symulator procesu dyfuzji informacji w sieciach złożonych, w oparciu o model kaskadowy (*independent cascade*).

Utworzony symulator jest zatem oparty na znanym modelu, którego użycie jest uzasadnione do badań nad planowaniem procesu rozprzestrzeniania informacji.

Istnieją alternatywne modele propagacji informacji w sieciach, które można by zastosować w tej sytuacji, a które nie były testowane przez autora – model liniowy (*linear threshold*) lub modele wywodzące się z epidemiologii (SIR lub SIS). Autor zdaje sobie sprawę z tych ograniczeń – w publikacji poświęconej symulatorowi omawia implementację alternatywnych modeli.

4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

Do najbardziej oryginalnych osiągnięć opisanych w rozprawie należą:

A. Postawienie i propozycja rozwiązania problemu planowania kampanii rozprzestrzeniania informacji w sieciach społecznych za pomocą metod wielokryterialnego wspomaganie decyzji. W publikacji A1 autor przedstawił przykładowy model decyzyjny oparty na trzech kryteriach parametrycznych kampanii (liczba węzłów inicjalnych kampanii, tzw. seeds, motywacja do przekazywania treści wyrażona jako średnie prawdopodobieństwo propagacji oraz miara centralności wykorzystywana do wyboru początkowych użytkowników sieci) oraz dwóch kryteriach wydajnościowych, wyliczanych na podstawie przeprowadzanych symulacji (czas wymagany na przeprowadzenie procesu rozprzestrzeniania informacji w sieci społecznej, wyrażony w liczbie iteracji oraz uzyskany zasięg w sieci).

B. innowacyjne badanie wywierania wpływu na dynamikę i zasięg procesu dyfuzji informacji poprzez oddziaływanie na rozkład prawdopodobieństwa propagacji w sieci – (*propagation probability spraying*). Oddziaływanie takie można osiągnąć dzięki stosowaniu zachęt, które zwiększają motywację wybranych uczestników sieci społecznej do przekazywania informacji podczas trwającej kampanii. Autor zwrócił uwagę, że koszt zachęt oferowanych różnym uczestnikom

w sieci społecznej może zależeć od ich centralności w tej sieci. Podobnie w realnych kampaniach koszt zaangażowania tzw. Influencerów zależy od ich centralności (zwykle wyrażanej stopniem

w sieci). W publikacji A6 opisane zostały badania nad wpływem kształtu rozkładu propagacji na proces propagacji informacji w sieci społecznej.

- C. Opracowanie zorientowanego obiektowo środowiska symulacyjnego do badania procesu dyfuzji informacji w sieciach złożonych. Autor zastosował paradygmat programowania obiektowego, by w języku R utworzyć narzędzie, które pozwala na badanie procesów rozprzestrzeniania informacji w sieciach złożonych w oparciu o model IC (ang. independent cascade). Za pomocą wymiennych modułów zasiewania, infekowania i logowania postępów środowisko to pozwala na badanie rozmaitych strategii kampanii rozprzestrzeniania informacji w sieciach rzeczywistych i syntetycznych. Symulator został udostępniony do wykorzystania przez innych członków społeczności naukowej badającej rozprzestrzenianie informacji w sieciach złożonych i opublikowany w artykule w czasopiśmie SoftwareX (200 pkt. ministerialnych) o otwartym dostępie (publikacja A10). Samo środowisko dostępne jest teraz publicznie na licencji GNU GPLv3.

5. Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienie uzyskanych przez siebie wyników (zwięzłość, jasność, poprawność redakcyjna rozprawy)?

Rozprawa jest napisana w języku polskim (wstęp) oraz angielskim (zbiór publikacji). Jest to dość dziwne i w moim odczuciu zmniejsza wartość rozprawy jako samodzielnej publikacji. Pomimo to, wstęp do rozprawy ma klarowną strukturę i jest dobrze napisany.

6. Jakie są słabe strony rozprawy i jej główne wady?

Najślabszym elementem rozprawy jest przegląd literatury. Rozprawa nie ma właściwie takiego rozdziału; literatura jest omawiana w rozdziale 4 wstępu, przy okazji omawiania poszczególnych zagadnień rozpatrywanych w różnych publikacjach wchodzących w skład rozprawy. Brak jest wniosków z przeglądu literatury i identyfikacji luk wiedzy, które wypełnia rozprawa.

Symulator opracowany w ramach badań autora nie został poddany walidacji. Można było się pokusić o taką walidację w oparciu o znane wyniki dotyczące propagacji informacji w dużych sieciach społecznych, na przykład wiadomości na Twitterze. Symulator oparty jest na wybranym modelu propagacji, modelu kaskadowym. Nie ma możliwości rozpatrywania innych modeli w celu weryfikacji rezultatów badań.

Wybrana przez autora metoda wspomaganie decyzji wielokryterialnych opiera się o preferencje wyrażane w postaci wag. Jak to zostało już wspomniane, jest to wybór niefortunny, albowiem taki



POLISH-JAPANESE ACADEMY OF INFORMATION TECHNOLOGY

ポ
ー
ラ
ン
ド
日
本
情
報
工
科
大
学

sposób wyrażania preferencji wymaga od decydenta dokonywania porównań parami. Utrudnia to znacznie wyrażenie preferencji – szczególnie w stosunku do bardzo różnych kryteriów, takich jak używane przez autora.

7. Jaka jest przydatność rozprawy dla Informatyki?

Informatyka jest nauką stosowaną, w której osiągnięciem badawczym jest opracowanie nowych metod przetwarzania informacji, które pozwalają rozwiązać realistyczne problemy o istotnym znaczeniu dla użytkowników. W swojej rozprawie doktorant zajmował się zagadnieniami opracowania wielokryterialnych metod wspomagania decyzji w celu planowania i ewaluacji procesów rozprzestrzeniania informacji w sieciach społecznych. Doktorant opracował szereg praktycznych modeli decyzyjnych, m.in. wykorzystujących próbki rzeczywistych sieci społecznych. W swojej pracy doktorant wielokrotnie kwestionował i modyfikował założenia swoich modeli, co świadczy

o krytycznym, naukowym podejściu do prowadzonych badań. Do osiągnięć doktoranta należy zaprojektowanie i implementacja zorientowanego obiektowo środowiska symulacyjnego do badania procesu dyfuzji informacji w sieciach złożonych, co również jest istotnym wkładem w dziedzinę Informatyki. Doktorant opracował też algorytm do wielokryterialnego doboru węzłów zasiewowych w celu targetowania wieloatrybutowych węzłów w procesie rozprzestrzeniania informacji w sieciach społecznych.



POLISH-JAPANESE ACADEMY OF INFORMATION TECHNOLOGY

ポ
ー
ラ
ン
ド
日
本
情
報
工
科
大
学

WNIOSKI

Rozprawa przedłożona przez mgr Artura Karczmarczyka stanowi zbiór publikacji. Publikacje, w których doktorant miał większościowy udział, osiągnęły łącznie ponad 700 punktów MNiSW według aktualnej punktacji. Świadczy to o bardzo dużym dorobku publikacyjnym doktoranta.

Rozprawa jest interesująca, jasno i klarownie zredagowana, i mimo pewnych mankamentów opisanych w recenzji, ma bardzo dużą wartość naukową.

Podsumowując stwierdzam, że **rozprawa spełnia wymagania stawiające rozprawom doktorskim przez obowiązujące przepisy z wyraźnym nadmiarem.**

Wniosuję o wyróżnienie rozprawy na podstawie wyraźnie wyróżniającego się dorobku publikacyjnego doktoranta.

prof. dr hab. Adam Wierzbicki

UCHWAŁA NR 145

**Senatu Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie
z dnia 27 czerwca 2022 r.**

**w sprawie nadania mgr. inż. Arturowi Karczmarczykowi
stopnia doktora**

Na podstawie art. 179 ust. 1 i 3 pkt 2b ustawy z dnia 3 lipca 2018 r. Przepisy wprowadzające ustawę – Prawo o szkolnictwie wyższym i nauce (Dz. U. poz. 1669, z późn. zm.) w związku z art. 14 ust. 2 pkt 5 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (tekst jedn. Dz. U. z 2017 r. poz. 1789, z późn. zm.) uchwała się, co następuje:

§ 1.

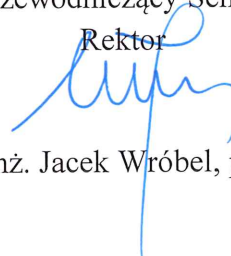
Senat Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie nadaje mgr. inż. Arturowi Karczmarczykowi stopień doktora w dziedzinie nauk inżynieryjno-technicznych, w dyscyplinie informatyka techniczna i telekomunikacja.

§ 2.

Uchwała wchodzi w życie z dniem podjęcia.

Przewodniczący Senatu

Rektor



dr hab. inż. Jacek Wróbel, prof. ZUT

UCHWAŁA NR 146
Senatu Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie
z dnia 27 czerwca 2022 r.

w sprawie wyróżnienia rozprawy doktorskiej dr. inż. Artura Karczmarczyka

Na podstawie art. 179 ust. 1 i 3 pkt 2b ustawy z dnia 3 lipca 2018 r. Przepisy wprowadzające ustawę – Prawo o szkolnictwie wyższym i nauce (Dz. U. poz. 1669, z późn. zm.) w związku z § 3 pkt 4 i § 21 rozporządzenia Prezesa Rady Ministrów z dnia 21 maja 2019 r. w sprawie kryteriów i trybu przyznawania nagród Prezesa Rady Ministrów oraz wzoru wniosku o ich przyznanie (Dz. U. poz. 976) oraz z § 11a uchwały nr 113 Senatu ZUT z dnia 23 września 2019 r. w sprawie określenia sposobu postępowania w sprawie nadania stopnia doktora w Zachodniopomorskim Uniwersytecie Technologicznym w Szczecinie (z późn. zm.) uchwała się, co następuje:

§ 1.

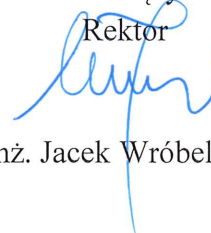
Senat Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie wyróżnia rozprawę doktorską dr. inż. Artura Karczmarczyka, pt. „Heterogeniczne oddziaływanie na procesy rozprzestrzeniania informacji w sieciach społecznych”.

§ 2.

Uchwała wchodzi w życie z dniem podjęcia.

Przewodniczący Senatu

Rektor



dr hab. inż. Jacek Wróbel, prof. ZUT